Dealing With Bias in Estimating Uncertainty and Risk

Stratis Gavaris

Dept. of Fisheries and Oceans, Marine Fish Division, Biological Station, St. Andrews, NB E0G 2X0, Canada.
E-mail address: GavarisS@mar.dfo-mpo.gc.ca

Abstract.- To quantify the uncertainty of fisheries stock assessment results or the risk of alternative management actions, we need to characterize the cumulative frequency distribution of the quantities of interest. Fisheries management quantities, such as biomass or fishing mortality rate, can only take positive values. Furthermore, estimators for fisheries management quantities from many assessment models are biased as a result of non-linearity in the models. Standard methods which assume a Gaussian distribution, failed to adequately account for the skew. The bootstrap percentile method did not adjust for the statistical estimation bias. The bootstrap bias corrected percentile technique appears to be best suited for general application.

Introduction

There is increasing recognition for the merits of explicitly taking into account the uncertainty of stock assessment results and the risks associated with alternate actions, when considering fisheries management decisions (Restrepo et al 1992). Incorporation of knowledge about uncertainty and risk for the provision of fisheries management advice is an integral aspect of the Precautionary Approach (ICES 1997). Practically, this implies that it is not sufficient to estimate the statistics for quantities of interest. We must also investigate their probability distributions.

The format in which uncertainty and risk results are conveyed is influenced by the specific management regime. In the Northeast USA, fishery managers are receiving this type of information in the form of cumulative frequency distributions and confidence intervals for the terminal year population quantities, such as spawning stock biomass (Anon. 1997). These uncertainties are also carried forward into short- and medium-term projections in order to evaluate alternative harvest strategies. In eastern Canada, fishery managers are receiving this kind of information in the form of a risk profile for achieving identified goals, such as an increase in spawning stock biomass, over a range of quota options in the forecast year (DFO 1997). They are concerned with the risk of achieving established reference points in the short-term projection if they choose a specific quota.

The risk profile is directly related to and derived from the cumulative frequency distributions of the estimated fisheries management quantities of interest. This is readily appreciated if one considers a surface constructed of cumulative frequency distributions for fishing mortality rate over a range of quotas (Fig. 1). The risk profile is the cross section of that surface at the established fishing mortality reference point. Therefore, the cumulative frequency distribution forms the basis of statements concerning uncertainty and risk.

It is recognised that estimators of fisheries management quantities from many typical fisheries assessment models are biased (Gavaris 1993, Prager 1994). This statistical bias arises from the non-linearity in the models. There may be other sources of bias, but here I only consider this statistical bias. The purpose of this study is to explore the impact of this statistical bias on the cumulative frequency distributions and the resulting risk profiles.

Methods

I describe three general methods for obtaining the cumulative frequency distribution of an arbitrary quantity, η, which is a function of estimated parameters, ξ, from some model. Let $\hat{\eta} = g(\hat{\xi})$ where g is the transformation function.

From an Assumed Distribution Type.

An obvious default method for constructing the cumulative frequency distribution is to invoke the Central Limit Theorem and assume that a Gaussian distribution adequately approximates the frequency distribution of the estimator $\hat{\eta}$. Applying the estimated statistics for $\hat{\eta}$, the desired cumulative frequency distribution is obtained by assuming $\eta \sim N(\hat{\eta} - Bias(\hat{\eta}), Var(\hat{\eta}))$. Note that this assumes the approximation $Var(\hat{\eta} - Bias(\hat{\eta})) \cong Var(\hat{\eta})$. This approach will be referred to as the standard method.

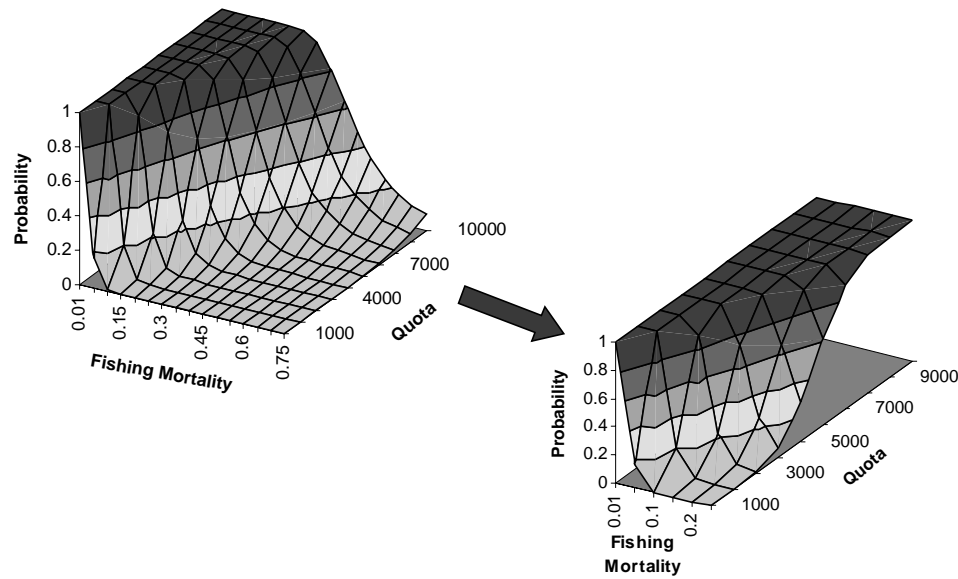The variance and bias of $\hat{\eta}$ can be obtained using the methods described in Ratkowsky (1983):

Figure 1. The risk that fishing mortality rate in the forecast year will exceed an established reference level, say 0.25, for some quota option, can be obtained as the cross section of the surface constructed from cumulative frequency distributions over that range of quotas.

$$Var(\hat{\eta}) = tr\left[GG^T \, \text{cov}(\hat{\xi}) \right]$$

$$Bias(\hat{\eta}) = G^T \, Bias(\hat{\xi}) + tr\left[W \, \text{cov}(\hat{\xi}) \right]/2 \ ,$$

where G is the vector of first derivatives of g with respect to parameters and W is the matrix of second derivatives of g with respect to parameters.

The covariance of the model parameters $\hat{\xi}$, can be estimated using the common linear approximation (Kennedy and Gentle 1980),

$$Cov(\hat{\xi}) = \hat{\sigma}^2 \left[J^T(\hat{\xi}) J(\hat{\xi}) \right]^{-1} \ ,$$

where $\hat{\sigma}^2$ is the mean square residual and $J(\hat{\xi})$ is the Jacobian matrix of the vector-valued objective function. The bias of the model parameters can be obtained using Box's (1971) approximation.

$$Bias(\hat{\xi}) =$$
$$\frac{-\hat{\sigma}^2}{2} \left(\sum_i J_i(\hat{\xi}) J_i^T(\hat{\xi}) \right)^{-1} \left(\sum_i J_i(\hat{\xi}) \right) tr\left[\left(\sum_i J_i(\hat{\xi}) J_i^T(\hat{\xi}) \right)^{-1} H_i(\hat{\xi}) \right],$$

where $J_i(\hat{\xi})$ are vectors of the first derivatives with respect to ξ of the vector-valued objective function and $H_i(\hat{\xi})$ are the Hessian matrices with respect to ξ.

From Bootstrapping.

Non-parametric bootstrap techniques offer the advantage of not making any assumptions about the error distribution. The bootstrap samples are used to calcu-

late the bootstrap replicate estimates, $\hat{\eta}^b$, of the quantity of interest. I considered two bootstrap methods, the percentile and the bias corrected percentile, for using the bootstrap replicate estimates to construct the cumulative frequency distribution.

The percentile method (Efron 1982) is a simple and direct way of forming an empirical cumulative frequency distribution. The probability that $\hat{\eta}$ is less than or equal to some value is defined as the proportion of bootstrap replicates, $\hat{\eta}^b$, less than or equal to that value:

$$\hat{\Omega}(x) = \text{Prob}\{\hat{\eta} \le x\} = \frac{\#\{\hat{\eta}^b \le x\}}{B} \ ,$$

where B is the total number of bootstrap replicates. For conceptual and graphing purposes, it is convenient to consider the empirical cumulative frequency distribution as the set of paired values $(\alpha, \bar{\eta}^b)$, where $\bar{\eta}^b$ are the ordered bootstrap replicates and α are the respective probability levels equal to $1/B, 2/B, 3/B, ... B/B$.

Frequently, the median of the bootstrap percentile density function does not equal the estimate obtained with the original data sample. The bias-corrected percentile method (Efron 1982) makes an adjustment for this type of bias. The bias-corrected percentile method can be thought of as an algorithm to replace the $\bar{\eta}^b$ in the paired values $(\alpha, \bar{\eta}^b)$ with the bias adjusted quantity $\bar{\eta}^b_{BC}$. The notation $\hat{\Omega}^{-1}(\ )$ or $\Phi^{-1}(\ )$ is used to represent the inverse distribution function, i.e. the critical value corresponding to the specified probability level. For each α in the paired values $(\alpha, \bar{\eta}^b)$, calculate the bias

adjusted quantity $\vec{\eta}_{BC}^{b}$:

$$\vec{\eta}_{BC}^{b} = \hat{\Omega}^{-1}\left(\Phi\left(2z_0 + z_\alpha\right)\right).$$

Here, Φ is the cumulative distribution function of a standard normal variate, $z_\alpha = \Phi^{-1}(\alpha)$ and $z_0 = \Phi^{-1}\left(\hat{\Omega}(\hat{\eta})\right)$. The term z_0 achieves the bias adjustment. If the median of the bootstrap density function is equal to $\hat{\eta}$, then $\hat{\Omega}(\hat{\eta})$ will be 0.5, z_0 will be zero, and $\vec{\eta}_{BC}^{b}$ will equal $\vec{\eta}^{b}$ (i.e. no bias adjustment). Note that computations are not carried out for $\alpha=B/B$ because $z_\alpha = \Phi^{-1}(\alpha = 1)$ is not defined.

Results

To illustrate the potential differences in outcomes, the three general techniques were applied to the results from a specific age structured analytical fisheries assessment model (Annex 1) using data from eastern Georges Bank haddock. In this example, the quantity of interest for fisheries management was spawning stock biomass, SSB. The cumulative frequency distributions of terminal year SSB, were derived and compared. The cumulative frequency distributions of SSB in the forecast year were also derived and used to obtain the risk of not achieving growth relative to the terminal year as a function of quota. In this example, a model-conditioned non-parametric bootstrap approach was employed. Bootstrap samples were obtained by adding the set of residuals obtained by sampling with replacement, to the model predicted values.

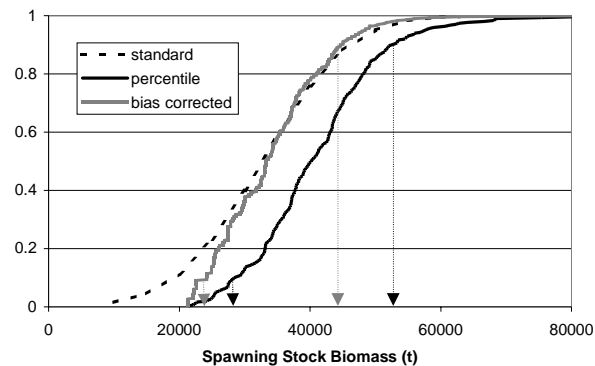Consider first the cumulative frequency distribution for SSB in the terminal year of the assessment, the type of advice portrayed in the NMFS SAW Advisory Reports. Figure 2 displays the results from the three approaches. The standard method gives the typically smooth and symmetric Gaussian distribution centred on the bias adjusted mean and characterised by the estimated variance. The empirical cumulative frequency distribution derived by the percentile method displays some skew and it is centred on the biased estimate. The corresponding 90% confidence interval associated with this approach gives values of [28,362t < SSB < 52,954t]. The bias-corrected percentile method appears to fully compensate for the bias and centres the empirical cumulative frequency distribution on the bias-adjusted estimate. It displays a greater degree of skew. For this example, the corresponding 90% confidence interval from the bias-corrected percentile method is more pessimistic with values of [24,249t < SSB < 44,968t].

Now consider how the statistical bias affects the risk profile for not achieving SSB growth in the forecast year, the kind of advice given in DFO Stock Status Reports. Recall that the risk profile is not a cumulative frequency distribution but a cross section of several cumulative frequency distributions. Figure 3 compares the results from the three approaches. Here again we see that the bias-corrected percentile method appears to compensate for the bias and results in a profile that is shifted towards that obtained with the standard method. The risk profile obtained from the percentile method gives a more optimistic outlook. For example, based on the percentile method, a 1997 quota of 5,000t implies an 18% risk of not achieving growth in SSB. This compares to a risk of 40% obtained from the bias-corrected risk profile.
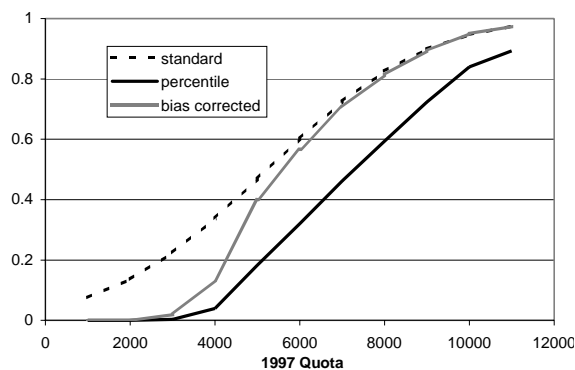

Figure 2. Comparison of the cumulative frequency distribution of spawning stock biomass in the terminal year indicates that the standard method does not reflect the skew displayed by the bootstrap methods. The percentile bootstrap method does not account for estimation bias and is shifted.

Figure 3. Comparison of the risk profiles for not achieving growth of the spawning stock biomass in the forecast year shows that results from the percentile method are shifted relative to the standard and bias corrected methods. For low risk levels, the distributional assumptions required by the standard method probably lead to erroneous results.

Discussion

Many fisheries management quantities of interest can only take positive values. In such instances, assuming a Gaussian approximation for these quantities does not capture the implied skew of their cumulative frequency distributions. This effect was apparent in the example using SSB. Consequently, the lower tail obtained with the standard method was considerably longer when compared to the bootstrap approaches. When the estimated variance is large, the lower confidence bound obtained from the standard method may be negative. It would appear that confidence statements based on results from the standard method might not be reliable for small cumulative probability levels. Assuming a lognormal distribution for some quantities may provide a better approximation, however theoretical justification may be lacking. For instance, in the example, it might be reasonable to assume, and there is some evidence to suggest, that the estimator of population abundance at age is lognormally distributed. The SSB then is the sum, multiplied by weight and maturity at age, of population abundance. The sum of lognormally distributed variables is not lognormal.

The bootstrap methods demonstrated that the empirical distribution for the SSB example was skewed. The results from the percentile method were shifted substantially, however. Confidence intervals or risk statements based on the percentile method can be markedly different from those based on the bias-corrected percentile method. Efron and Tibshirani (1993) argue that confidence statements based on the bias-corrected and accelerated method offer a substantial improvement over the percentile method, both in theory and in practice. The accelerated method was not used here. Loh and Wu (1987) indicated that the accelerated method might offer only marginal improvement over the bias-corrected method, but this aspect merits further investigation for the stock assessment problem. Nevertheless, we may conclude that the bias-corrected method should provide more accurate confidence statements than does the simple percentile method.

Recognising the potential to inadequately characterise the shape of the frequency distribution with the standard method, and the failure of the percentile method to account for estimation bias, the bootstrap bias-corrected percentile technique is recommended for general application.

Literature Cited

Anonymous. 1997. 24ᵗʰ Northeast Regional Stock Assessment Workshop Advisory Report on Stock Status. Available from NOAA, National Marine Fisheries Service, North-east Fisheries Science Center, Woods Hole, MA 02543-1097.

Box, M.J. 1971. Bias in nonlinear estimation. J.R. Statist. Soc., Ser. B 33: 171-201.

DFO. 1997. Eastern Georges Bank Haddock. DFO Sci. Stock Status Report A3-08. available from Maritimes Regional Advisory Process, Dept. of Fisheries and Oceans, P.O. Box 1006, Stn. B105, Dartmouth, NS, Canada B2Y 4A2 or from the internet at http://www.dfo-mpo.gc.ca/csas/

Efron, B. 1982. The jacknife, the bootstrap and other resampling plans. CBMS Monograpgh #38, Society for Industrial and Applied Mathematics, Philadelphia.

Efron, B. and R.J. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall. New York. 436pp.

Gavaris, S. 1993. Analytical estimates of reliability for the projected yield from commercial fisheries, p. 185-191. In S. J. Smith, J. J. Hunt and D. Rivard (ed.) Risk evaluation and biological reference points for fisheries management. Can. Spec. Publ. Fish. Aquat. Sci. 120.

ICES. 1997a. Report of the comprehensive fishery evaluation working group. ICES CM 1997/Assess:15. 140pp.

Kennedy, W.J., Jr. and J.E. Gentle. 1980. Statistical computing. Marcel Dekker. New York. 591 pp.

Loh, W. and C.F.J. Wu. 1987. Comment on Better bootstrap confidence intervals. J. Amer. Statist. Assoc. 82(397):188-190.

Prager, M.H. 1994. A suite of extensions to a nonequilibrium surplus-production model. Fish. Bull. (U.S.) 92: 374-389.

Ratkowsky, D.A. 1983. Nonlinear regression modeling. Marcel Dekker. New York. 276 pp.

Restrepo,V.R., J.M.Hoenig, J.E. Powers, J.W. Baird and S.C. Turner. 1992. A simple simulation approach to risk and cost analysis, with applications to swordfish and cod fisheries. Fish. Bull. U.S. 90: 736-748.

Annex 1: Fisheries Assessment Model

The available data were:

$C_{a,t}$ = catch at age, age $a = 0, 1 \ldots 8$, time $t = 1986$, 1987 ... 1996.

$I_{a,t}$ = DFO spring survey, age $a = 1, 2 \ldots 8$, time $t = 1986$, 1987... 1997.

The employed model formulation assumed that the error in the catch at age was negligible. The errors in the abundance indices were assumed independent and identically distributed after taking natural logarithms of the values. The annual natural mortality rate, M, was assumed constant and equal to 0.2. A model formulation using as parameters the natural logarithm of population abundance at the beginning of the year was considered because of close-to-linear behavior for such a parameterization (Gavaris 1993). Thus, a total of 16 parameters were estimated:

$\theta_{a,t'}$, ages $a = 1, 2, \ldots 8$ at time t' 1997,
κ_a, ages $a = 1, 2, \ldots 8$.

A solution for the parameters was obtained by minimizing the sum of squared differences between the natural logarithm observed abundance indices and the natural logarithm population abundance adjusted for catchability by the calibration constants:

$$\Psi\left(\hat{\theta}, \hat{\kappa}\right) = \sum_{a,t} \left(\psi_{a,t}\left(\hat{\theta}, \hat{\kappa}\right)\right)^2 = \sum_{a,t} \left(\ln I_{a,t} - \left(\hat{\kappa}_a + \ln N_{a,t}\left(\hat{\theta}\right)\right)\right)^2$$

For convenience, the model's population abundance $N_{a,t}\left(\hat{\theta}\right)$ is abbreviated by $N_{a,t}$. At time t', the population abundance was obtained directly from the parameter estimates, $N_{a,t'} = e^{\hat{\theta}_{a,t'}}$. For all other times, the population abundance was computed using the virtual population analysis algorithm, which incorporates the common exponential decay model

$$N_{a+\Delta t, t+\Delta t} = N_{a,t} e^{-\left(F_{a,t} + M_a\right)\Delta t} .$$

Year was used as the unit of time. Therefore, ages were expressed as years and the fishing and natural mortality rates were annual instantaneous rates. The fishing mortality rate, $F_{a,t}$, exerted during the time interval t to $t + \Delta t$, was obtained by solving the catch equation,

$$C_{a,t} = \frac{F_{a,t} \Delta t N_{a,t} \left(1 - e^{-\left(F_{a,t} + M_a\right)\Delta t}\right)}{\left(F_{a,t} + M_a\right)\Delta t} ,$$

using a Newton-Raphson algorithm. The fishing mortality rate for the oldest age in the last time interval of each year was assumed equal to the weighted average for ages fully recruited to the fishery during that time interval

$$F_{8,t} = \sum_{a=4}^{7} N_{a,t} F_{a,t} \Bigg/ \sum_{a=4}^{7} N_{a,t} .$$